

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**



PCT/AU 96 / 00094

REC'D 08 MAR 1996  
WIPO PCT

Patents Office  
Canberra

I, DAVID DANIEL CLARKE, ASSISTANT DIRECTOR PATENT SERVICES, hereby certify that the annexed are true copies of the Provisional specification and drawing(s) as filed on 20 November 1995 in connection with Application No. PN 6647 for a patent by THE COUNCIL OF THE QUEENSLAND INSTITUTE OF MEDICAL RESEARCH and AMRAD CORPORATION LIMITED filed on 20 November 1995.

I further certify that the annexed documents are not, as yet, open to public inspection.

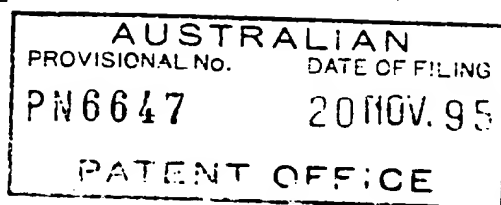
PRIORITY DOCUMENT



WITNESS my hand this Twenty-eighth  
day of February 1996

A handwritten signature in cursive script, appearing to read "D Clarke".

DAVID DANIEL CLARKE  
ASSISTANT DIRECTOR PATENT SERVICES



THE COUNCIL OF THE  
QUEENSLAND INSTITUTE OF  
MEDICAL RESEARCH and  
AMRAD CORPORATION LIMITED

A U S T R A L I A  
Patents Act 1990

PROVISIONAL SPECIFICATION  
for the invention entitled:

"A NOVEL GROWTH FACTOR AND A GENETIC SEQUENCE ENCODING SAME"

The invention is described in the following statement:

## A NOVEL GROWTH FACTOR AND A GENETIC SEQUENCE ENCODING SAME

5 The present invention relates generally to an isolated molecule having vascular endothelial growth factor-like properties and to a genetic sequence encoding same. The molecule will be useful in the development of a range of therapeutics and diagnostics useful in the treatment, prophylaxis and/or diagnosis of conditions requiring enhanced or diminished vasculature and/or vascular permeability.

10

Bibliographic details of the publications referred to by author in this specification are collected at the end of the description. Sequence Identity Numbers (SEQ ID NOs.) for the nucleotide and amino acid sequences referred to in the specification are defined following the bibliography.

15

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element or integer or group of elements or integers but not the exclusion of any other element or integer or group of elements or integers.

20

Vascular endothelial growth factor (hereinafter referred to as "VEGF"), also known as vasoactive permeability factor, is a secreted, covalently linked homodimeric glycoprotein that specifically activates endothelial tissues (Senger *et al.*, 1993). A range of functions have been attributed to VEGF such as its involvement in normal angiogenesis including  
25 formation of the corpus luteum (Yan *et al.*, 1993) and placental development (Sharkey *et al.*, 1993), regulation of vascular permeability (Senger *et al.*, 1993), inflammatory angiogenesis (Sunderkotter *et al.*, 1994) and autotransplantation (Dissen *et al.*, 1994) and human diseases such as tumour promoting angiogenesis (Folkman & Shing, 1992), rheumatoid arthritis (Koch *et al.*, 1994) and diabetes related retinopathy (Folkman &  
30 Shing, 1992).

VEGF is, therefore, an important molecule making it a potentially valuable target for research into therapeutics, prophylactics and diagnostic agents based on VEGF or its activities. There is also a need to identify homologues or otherwise related molecules  
5 for use as an alternative to VEGF or in conjunction with VEGF.

In work leading up to the present invention, the inventors sought the multiple endocrine neoplasia type I susceptibility gene (MEN1). Surprisingly, the inventors discovered that a genetic sequence excluded as a candidate for the MEN1 gene was nevertheless a new  
10 growth factor having some similarity to VEGF.

Accordingly, one aspect of the present invention comprises a biologically isolated proteinaceous molecule comprising a sequence of amino acids which:

- (i) is at least about 15% similar to the amino acid sequence set forth in SEQ ID  
15 NO:2; and
- (ii) is at least 5% dissimilar to the amino acid sequence set forth in SEQ ID NO:2.

Another aspect of the present invention provides a biologically isolated proteinaceous molecule having the following characteristics:

- 20 (i) comprises an amino acid sequence having at least about 15% similarity but at least about 5% dissimilarity to all or part of the amino acid sequence set forth in SEQ ID NO:2;
- (ii) exhibits at least one property in common with VEGF.

25 A related aspect of the present invention contemplates a biologically isolated proteinaceous molecule having the following characteristics:

- (i) comprises an amino acid sequence having at least about 15% similarity but at least about 5% dissimilarity to the amino acid sequence set forth in SEQ ID NO:2;
- 30 (ii) exhibits at least one of the following properties:
  - (a) ability to induce proliferation of vascular endothelial cells;
  - (b) ability to interact with *flt-1/flk-1* family of receptors;

- (c) ability to induce cell migration, cell survival and/or an increase in intracellular levels of alkaline phosphatase.

By "biologically isolated" is meant that the molecule has undergone at least one step of purification from a biological source. Preferably, the molecule is also biologically pure meaning that a composition comprises at least about 20%, more preferably at least about 40%, still more preferably at least about 65%, even still more preferably at least about 80-90% or greater of the molecule as determined by weight, activity or other convenient means, relative to other compounds in the composition. Most preferably, the molecule is sequencably pure.

Another preferred aspect of the present invention provides the molecule in recombinant form.

According to this aspect of the present invention, there is provided a recombinant molecule comprising a sequence of amino acids which:

- (i) is at least about 15% similar to the amino acid sequence set forth in SEQ ID NO:2; and
- (ii) is at least 5% dissimilar to the amino acid sequence set forth in SEQ ID NO:2.

A related aspect of the present invention is directed to a recombinant molecule having the following characteristics:

- (i) comprises an amino acid sequence having at least about 15% similarity but at least about 5% dissimilarity to all or part of the amino acid sequence set forth in SEQ ID NO:2;
- (ii) exhibits at least one property in common with VEGF.

A further related aspect of the present invention contemplates a recombinant molecule having the following characteristics:

- (i) comprises an amino acid sequence having at least about 15% similarity but at least about 5% dissimilarity to the amino acid sequence set forth in SEQ ID NO:2;

- (ii) exhibits at least one of the following properties:
- (a) ability to induce proliferation of vascular endothelial cells;
  - (b) ability to interact with *flt-1/flk-1* family of receptors;
  - (c) ability to induce cell migration, cell survival and/or an increase in
- 5 intracellular levels of alkaline phosphatase.

The present invention also contemplates genomic or partial genome clones encoding a proteinaceous molecule having at least about 15% amino acid similarity but at least about 5% dissimilarity to SEQ ID NO:1.

10

The amino acid sequence set forth in SEQ ID NO:2 corresponds to human VEGF (referred to herein as "VEGF<sub>165</sub>"). Accordingly, the molecule of the present invention is VEGF-like or is a homologue of VEGF but comprises an amino acid sequence which is similar but non-identical to the amino sequence of VEGF. Although the present

15 invention is exemplified using a human VEGF-like molecule, this is done with the understanding that the instant invention contemplates the homologous molecule and encoding sequence from other mammals such as livestock animals (e.g. sheep, pigs, horses and cows), companion animals (e.g. dogs and cats) and laboratory test animals (e.g. mice, rats, rabbits and guinea pigs) as well as non-mammals such as birds (e.g.

20 poultry birds), fish and reptiles. In a most preferred embodiment, the VEGF-like molecule is of human origin and encoded by a gene located at chromosome 11q13. The present invention extends, therefore, to the genomic sequence or part thereof encoding the subject VEGF-like molecule.

25 Preferably, the percentage similarity is at least about 30%, more preferably at least about 40%, still more preferably at least about 50%, still even more preferably at least about 60-70%, yet even more preferably at least about 80-95% to all or part of the amino acid sequence set forth in SEQ ID NO:2.

30 In a particularly preferred embodiment, the VEGF-like molecule of the present invention comprises a sequence of amino acids as set forth in SEQ ID NO:4 or is a part, fragment, derivative or analogue thereof. The amino acid sequence set forth in SEQ ID NO:4 is



also referred to herein as "SOM175<sub>short</sub>". Particularly preferred similarities include about 19-20%, and 29-30%. Reference herein to derivatives also includes splice variants. Accordingly, the present invention extends to splice variants of SOM175<sub>short</sub>. Examples of splice variants contemplated by the present invention include but are not  
5 limited to variants with an amino acid sequence substantially as set forth in at least one of SEQ ID NO:6, SEQ ID NO:8 and/or SEQ ID NO:10 or mutants or derivatives or further splice variants thereof.

Another embodiment provides a recombinant molecule having the following  
10 characteristics:

- (i) an amino acid sequence substantially as set forth in SEQ ID NO:4 or having at least about 15% similarity to all or part thereof provided that said amino acid sequence is at least about 5% dissimilar to all or part of the amino acid sequence set forth in SEQ ID NO:2;
- 15 (ii) exhibits at least one biological property in common with VEGF.

Another embodiment provides a recombinant molecule having the following characteristics:

- 20 (i) an amino acid sequence substantially as set forth in SEQ ID NO:6 or having at least about 15% similarity to all or part thereof provided that said amino acid sequence is at least about 5% dissimilar to all or part of the amino acid sequence set forth in SEQ ID NO:2;
- (ii) exhibits at least one biological property in common with VEGF.

25 Another embodiment provides a recombinant molecule having the following characteristics:

- (i) an amino acid sequence substantially as set forth in SEQ ID NO:8 or having at least about 15% similarity to all or part thereof provided that said amino acid sequence is at least about 5% dissimilar to all or part of the amino acid sequence set forth in SEQ ID NO:2;
- 30 (ii) exhibits at least one biological property in common with VEGF.

Another embodiment provides a recombinant molecule having the following characteristics:

- (i) an amino acid sequence substantially as set forth in SEQ ID NO:10 or having at least about 15% similarity to all or part thereof provided that said amino acid sequence is at least about 5% dissimilar to all or part of the amino acid sequence set forth in SEQ ID NO:2;
- (ii) exhibits at least one biological property in common with VEGF.

Such properties of VEGF include at least one of:

- (a) ability to induce proliferation of vascular endothelial cells;
  - (b) an ability to interact with *flt-1/flk-1* family of receptors;
  - (c) an ability to induce cell migration, cell survival and/or an increase in intracellular levels of alkaline phosphatase.
- 15 In accordance with the present invention, a preferred similarity is at least about 40%, more preferably at least about 50% and even more preferably at least about 65% similarity.

20 Still a further aspect of the present invention contemplates a peptide fragment corresponding to a portion of the amino acid sequence set forth in SEQ ID NO:4 or a splice variant thereof such as set forth in SEQ ID NO:6, SEQ ID NO:8 or SEQ ID NO:10 or a chemical equivalent thereof. The biologically isolated or recombinant molecule of the present invention may be naturally glycosylated or may comprise an altered glycosylation pattern depending on the cells from which it is isolated or  
25 synthesised. For example, if produced by recombinant means in prokaryotic organisms, the molecule would be non-glycosylated. The molecule may be a full length, naturally occurring form or may be a truncated or otherwise derivatised form.

30 Yet another aspect of the present invention is directed to a nucleic acid molecule encoding the VEGF-like molecule herein described. More particularly, the present invention provides a nucleic acid molecule comprising a sequence of nucleotides substantially as set forth in SEQ ID NO:3 or having at least 15% similarity to all or part

thereof or being capable of hybridising under low stringency conditions to a reverse complement of the nucleotide sequence as set forth in SEQ ID NO:3 provided that the nucleic acid sequence having at least 15% similarity but at least 30% dissimilarity to the nucleotide sequence as set forth in SEQ ID NO:3. The nucleotide sequence set forth in  
5 SEQ ID NO:3 is also referred to herein as "SOM175". Preferably, the percentage dissimilarity is about 35%, more preferably about 39% and even more preferably about 40-50% or greater.

For the purposes of defining the level of stringency, reference can conveniently be made  
10 to Sambrook *et al* (1989) at pages 9.47-9.51 which is herein incorporated by reference where the washing steps disclosed are considered high stringency. A low stringency is defined herein as being in 4-6X SSC/0.1-0.5% w/v SDS at 37-45°C for 2-3 hours. Depending on the source and concentration of nucleic acid involved in the hybridisation, alternative conditions of stringency may be employed such as medium stringent  
15 conditions which are considered herein to be 1-4X SSC/0.25-0.5% w/v SDS at  $\geq 45^{\circ}\text{C}$  for 2-3 hours or high stringent conditions considered herein to be 0.1-1X SSC/0.1% w/v SDS at  $60^{\circ}\text{C}$  for 1-3 hours.

The present invention further contemplates a nucleic acid molecule which encodes a  
20 VEGF-like molecule as hereinbefore described having at least 15% nucleotide sequence homology to SEQ ID NO:3. Preferred levels of homology include at least about 40%, more preferably around 60-70%.

The VEGF-like molecule of the present invention will be useful in the development of  
25 a range of therapeutic and/or diagnostic applications alone or in combination with other molecules such as VEGF. The present invention extends, therefore, to pharmaceutical compositions comprising the VEGF-like molecule or parts, fragments, derivatives, homologues or analogues thereof together with one or more pharmaceutically acceptable carriers and/or diluents. Furthermore, the present invention extends to vectors  
30 comprising the nucleic acid sequence set forth in SEQ ID NO:3 or having at least about 15%, more preferably about 40% and even more preferably around 60-70% similarity thereto but at least 30% and more preferably around 39% dissimilarity thereto and host

cells comprising same. In addition, the present invention extends to ribozymes and antisense molecules based on SEQ ID NO:3 as well as neutralizing antibodies to the VEGF-like molecule. Such molecules may be useful in ameliorating the effects of, for example, over expression of VEGF-like genes leading to angiogenesis or vascularization  
5 of tumours.

The present invention also contemplates antibodies to the VEGF-like molecule or nucleic acid probes to a gene encoding the VEGF-like molecule which are useful as diagnostic agents.

10

The present invention is further described by reference to the following non-limiting Figures and/or Examples.

In the Figures:

15

**Figure 1** Nucleotide sequence [SEQ ID NO:1] and corresponding amino acid sequence [SEQ ID NO:2] of VEGF<sub>165</sub>.

**Figure 2** Nucleotide sequence [SEQ ID NO:3] and corresponding amino acid  
20 sequence [SEQ ID NO:4] of SOM175.

**Figure 3** Results of BLAST search with SOM175 protein sequence.

**Figure 4** BESTFIT alignment of VEGF cDNA and SOM175 cDNA.

25

**Figure 5** Multiple alignment of VEGF<sub>165</sub> with SOM175 and its splice variants at the nucleotide level.

**Figure 6** Multiple alignment of VEGF<sub>165</sub> with SOM175 and its splice variants at  
30 the amino acid level.

**Figure 7** Diagrammatic representation of SOM175 and its splice variants.

**Figure 8(a)** Diagrammatic representation of genomic structure of human SOM175 genomic showing exon/intron map.

**Figure 8(b)** Diagrammatic representation of genomic structure of human SOM175 showing exon/intron boundaries.

TABLE 1  
SUMMARY OF SEQUENCE IDENTITY NUMBERS

10

SEQ ID NO:1	Nucleotide sequence of VEGF <sub>165</sub>
SEQ ID NO:2	Amino acid sequence of VEGF <sub>165</sub>
SEQ ID NO:3	Nucleotide sequence of SOM175 (VEGF-like molecules)
SEQ ID NO:4	Amino acid sequence of SOM175
15 SEQ ID NO:5	Nucleotide sequence of SOM175 absent exon 6
SEQ ID NO:6	Amino acid sequence of SOM175 absent exon 6
SEQ ID NO:7	Nucleotide sequence of SOM175 absent exon 6 and exon 7
SEQ ID NO:8	Amino acid sequence of SOM175 absent exon 6 and exon 7
SEQ ID NO:9	Nucleotide sequence of SOM175 absent exon 4
20 SEQ ID NO:10	Amino acid sequence of SOM175 absent exon 4

## EXAMPLE 1

### Human cDNA clones

The original SOM175 cDNA was isolated by screening a human foetal brain library  
5 (λzapII, Stratagene) with the cosmid D11S750 (Larsson *et al*, 1992) using methods  
previously described (Xu *et al* 1992). The plasmid was excised "*in vivo*" and a single  
1.1kb cDNA was obtained. Three independent SOM175 cDNAs clones were also  
isolated from a human foetal spleen library (Stratagene, Uni-zap) using the above-  
mentioned SOM175 insert as a probe. Three clones were obtained: SOM175-4A, -5A  
10 and -6A. SOM175-5A is an alternately spliced clone with exon 4 being absent  
(SOM175-e4). These library screens were performed using hybridisation conditions  
recommended by the manufacturer of the library (Stratagene) and random primed insert  
of SOM175.

15 Two partial human SOM175 cDNAs have also isolated from a λGT11 human melanoma  
cell line A2058 library (Clontech) cDNA library screens were performed using  
hybridisation conditions described by Church and Gilbert, 1984). In each case, the  
probe was generated by random priming of a PCR product derived from SOM175 (18f-  
700r).

20

### Mouse cDNA Clones

Human SOM175 was also used to screen a mouse neonatal whole brain cDNA library  
(Unizap, Stratagene). Four non-chimeric clones were isolated: M175-A, B, C, D. All  
clones were partial cDNAs and M175-C contained several introns. Three of these  
25 cDNAs lacked the exon 6.

Another clone referred to as M1 was completely sequenced and was found to contain  
the full open reading frame plus part of the 5'utr and total 3'utr.

## EXAMPLE 2

### DNA SEQUENCE ANALYSIS

5 The entire sequence of the cDNA clone (SOM175) was compiled and is shown in Figure 2 with its corresponding amino acid sequence. This sequence was screened for open reading frames using the MAP program (GCG, University of Wisconsin). A single open reading frame of 672bp was observed (see Figure 2). There appears to be little 5' untranslated sequences (2bp). The 3' untranslated region appears to be complete as it includes a poly-adenylation signal and poly-A tail.

10

Database homology searches were performed using the BLAST algorithm (run at NCBI, USA). This analysis revealed homology to several mammalian forms of VEGF (see Figure 3). The amount of homology between SOM175 and human VEGF<sub>165</sub> was determined using the BESTFIT program (GCG, University of Wisconsin; see Figures 15 4 and 5). Nucleotide homology was estimated at 69.7% and protein homology was estimated as at least 33.3% identity and 52.5% conservation using BESTFIT analysis. BLAST analysis on nucleotide sequences revealed the almost complete match to a human expressed sequence tag EST06302 (Adams *et al.*, 1993).

20 These data indicate that SOM175 encodes a growth factor that has structural similarities to VEGF. Both genes show start and stop codons in similar positions and share discrete blocks of homology. All 8 cysteines as well as a number of other VEGF residues believed to be involved in dimerisation are conserved. These residues are Cysteine-47, Proline-70, Cysteine-72, Valine-74, Arginine-77, Cysteine-78, Glycine-80, Cysteine-81, 25 Cysteine-82, Cysteine-89, Proline-91, Cysteine-122 and Cysteine-124 and are shown in Figure 6. Given the structural conservation between VEGF and the SOM175 gene product it is also possible that they share functional similarities. It is proposed that SOM175 encodes a VEGF-like molecule that shares some properties with VEGF but has unique properties of its own. The nucleotide sequence and corresponding amino acid 30 sequence of VEGF<sub>165</sub> is shown in Figure 1.

### EXAMPLE 3

The percentage similarity and divergence between VEGF<sub>165</sub> family and SOM175 family (protein) were analysed using the Clustal method, MegAlign Software, DNASTAR, Wisconsin. The results are shown in Tables 2.1 and 2.2. The alternatively spliced forms of SOM175 are abbreviated to SOM175-e6 where all of exon 6 is deleted; SOM175-e6 and 7 where all of exons 6 and 7 are deleted; and SOM175-e4 where all of exon 4 is deleted. The spliced form of SOM175 are shown in Figure 7. Genomic maps of SOM175 showing intron/exon boundaries are shown in Figure 8a and 8b.

Table 2.1

A Percent nucleotide similarity between splice variants of SOM175 and human VEGF<sub>165</sub>

	VEGF <sub>165</sub>	SOM175	SOM175-e6	SOM175-e6&7	SOM175-e4
VEGF <sub>165</sub>	***	34.9	39.7	41.4	37.0
SOM175		***	98.9	95.1	99.2
SOM175-e6			***	98.8	84.0
SOM175-e6&7				***	80.3
SOM175-e4					***



**B Percent nucleotide divergence between splice variants of SOM175 and human VEGF<sub>165</sub>**

5		VEGF <sub>165</sub>	SOM175	SOM175-e6	SOM175-e6&7	SOM175-e4
	VEGF <sub>165</sub>	***	41.7	41.6	41.7	41.8
	SOM175		***	0.2	0.2	0.0
	SOM175-e6			***	0.0	0.2
10	SOM175-e6&7				***	0.3
	SOM175-e4					***

**Table 2.2**

**15 A Percent amino acid identity between splice variants of SOM175 and human VEGF<sub>165</sub>**

		VEGF <sub>165</sub>	SOM175	SOM175-e6	SOM175-e6&7	SOM175-e4
20	VEGF <sub>165</sub>	***	31.4	42.3	33.5	40.6
	SOM175		***	74.7	73.7	99.1
	SOM175-e6			***	76.8	99.1
	SOM175-e6&7				***	99.1
	SOM175-e4					***
25						

**B Percent amino acid divergence between splice variants of SOM175 and human VEGF<sub>165</sub>**

5		VEGF <sub>165</sub>	SOM175	SOM175-e6	SOM175-e6&7	SOM175-e4
	VEGF <sub>165</sub>	***	65.7	55.4	54.6	57.4
	SOM175		***	19.9	4.2	0.0
	SOM175-e6			***	0.0	0.0
10	SOM175-e6&7				***	0.0
	SOM175-e4					***

15

**EXAMPLE 4**

**BIOASSAYS TO DETERMINE THE FUNCTION OF SOM175**

Assays are conducted to evaluate whether SOM175 has similar activities to VEGF on endothelial cell function, angiogenesis and wound healing. Other assays are performed based on the results of receptor binding distribution studies.

20

**Assays of endothelial cell function**

*Endothelial cell proliferation.* Endothelial cell growth assays as described in Ferrara & Henzel (1989) and in Gospodarowicz *et al* (1989).

25

*Vascular permeability assay.* This assay, which utilises the Miles test in guinea pigs, will be performed as described in Miles & Miles (1952).

*Cell adhesion assay.* The influence of SOM175 on adhesion of polymorphs to endothelial cells is analysed.

30

*Chemotaxis.* This is performed using the standard Boyden chamber chemotaxis assay.

*Plasminogen activator assay.* Endothelial cells are tested for plasminogen activator and plasminogen activator inhibitor production upon addition of SOM175 (Pepper *et al* (1991)).

- 5 *Endothelial cell migration assay.* The ability of SOM175 to stimulate endothelial cells to migrate and form tubes is assayed as described in Montesano *et al* (1986).

#### **Angiogenesis Assay**

- SOM175 induction of an angiogenic response in chick chorioallantoic membrane is  
10 evaluated as described in Leung *et al* (1989).

Possible neurotrophic actions of SOM175 are assessed using the following assays:

#### **Neurite outgrowth assay and gene induction (PC12 cells)**

- 15 PC12 cells (a phaeochromocytoma cell line) respond to NGF and other neurotrophic factors by developing the characteristics of sympathetic neurons, including the induction of early and late genes and the extension of neurites. These cells are exposed to SOM175 and their response monitored (Drinkwater *et al* (1991); and Drinkwater *et al* (1993)).

20

#### **Cultured neurons from the Peripheral Nervous System (PNS)**

Primary cultures of the following PNS neurons are exposed to SOM175 and monitored for any response:

- sensory neurons from neural crest and dorsal root ganglia
- 25 - sympathetic neurons from sympathetic chain ganglia
- placode derived sensory neurons from nodose ganglia
- motoneurons from spinal cord

The assays are described in Suter *et al* (1992) and in Marinou *et al* (1992).

- 30 Where an *in vitro* response is observed, *in vivo* assays for properties such as uptake and retrograde transport are performed as described in Hendry *et al* (1992).

### **Nerve regeneration (PNS)**

Where neurotrophic effects of SOM175 are observed, its possible role in the regeneration of axotomised sensory neurons, sympathetic neurons and motoneurons is analysed by the methods of Otto *et al* (1989); Yip *et al* (1984) and Hendry *et al* 5 (1976).

### **Actions of SOM175 on CNS neurons**

The ability of SOM175 to promote survival of central nervous system neurons is analysed as described in Hagg *et al* (1992); Williams *et al* (1986); Hefti (1986) and 10 Kromer (1987).

### **Wound Healing**

The ability of SOM175 to support wound healing are tested in the most clinically relevant model available, as described in Schilling *et al* (1959) and utilised by Hunt 15 *et al* (1967).

### **The Haemopoietic System**

A variety of *in vitro* and *in vivo* assays on specific cell populations of the haemopoietic system are available and are outlined below:

#### **20 *Stem Cells***

##### ***Murine***

A variety of novel *in vitro* murine stem cell assays have been developed using FACS-purified cells:

#### **(a) Repopulating Stem Cells**

25 These are cells capable of repopulating the bone marrow of lethally irradiated mice, and have the Lin<sup>-</sup>, Rh<sup>hi</sup>, Ly-6A/E<sup>+</sup>, c-kit<sup>+</sup> phenotype. The test substance is tested on these cells either alone, or by co-incubation with multiple factors, followed by measurement of cellular proliferation by <sup>3</sup>H thymidine incorporation.

**(b) Late Stage Stem Cells**

- These are cells that have comparatively little bone marrow repopulating ability but can generate D13 CFU-S. These cells have the Lin<sup>-</sup>, Rh<sup>hi</sup>, Ly-6A/E<sup>+</sup>, c-kit<sup>+</sup> phenotype. The test substance is incubated with these cells for a period of time, injected into lethally irradiated recipients, and the number of D13 spleen colonies enumerated.

**(c) Progenitor-Enriched Cells**

- These are cells that respond *in vitro* to single growth factors, and have the Lin<sup>-</sup>, Rh<sup>hi</sup>, Ly-6A/E<sup>+</sup>, c-kit<sup>+</sup> phenotype. This assay will show if SOM175 can act directly on haemopoietic progenitor cells. The test substance is incubated with these cells in agar cultures, and the number of colonies enumerated after 7-14 days.

**15 Atherosclerosis**

- Smooth muscle cells play a crucial role in the development or initiation of atherosclerosis, requiring a change in their phenotype from a contractile to a synthetic state. Macrophages, endothelial cells, T lymphocytes and platelets all play a role in the development of atherosclerotic plaques by influencing the growth and phenotypic modulations of smooth muscle cell. An *in vitro* assay that measures the proliferative rate and phenotypic modulations of smooth muscle cells in a multicellular environment is used to assess the effect of SOM175 on smooth muscle cells. The system uses a modified Rose chamber in which different cell types are seeded onto opposite coverslips.

25

**Effects of SOM175 on bone**

- The ability of SOM175 to regulate proliferation of osteoblasts is assayed as described in Lowe *et al* (1991). Any effects on bone resorption are assayed as described in Lowe *et al* (1991). Effects on osteoblast migration and changes in intracellular molecules (e.g. cAMP accumulation, alkaline phosphatase levels) are analysed as described in Midy *et al* (1994).

**Effects on skeletal muscle cells**

Effects of SOM175 on proliferation of myoblasts and development of myotubes can be determined as described by Ewton *et al* (1980) and by Gospodarowicz *et al* (1976).

5

Those skilled in the art will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications.

The invention also includes all of the steps, features, compositions and compounds  
10 referred to or indicated in this specification, individually or collectively, and any and all combinations of any two or more of said steps or features.

## BIBLIOGRAPHY

- Adams MD, Soares MB, Kerlavage AR, Fields C, Venter JC, (1993) *Nature Genet.*, **4**, 373-380.
- Church and Gilbert, 1984.**
- Dissen GA, Lara HE, Fabrenbach WH, Costa ME, Ojeda SR, (1994) *Endocrinology* **134**, 1146-1154.
- Drinkwater CC, Barker PA, Suter U and Shooter EM (1993) *J. Biol. Chem.*, **268**, 23202-23207.
- Drinkwater CC, Suter U, Angst C and Shooter EM (1991) *Proc. Roy. Soc. Lond. (Series B)*, **246**, 307-313.
- Ewton DZ & Florini JR (1980) *Endocrinology*, **106**: 577-583.
- Ferrara N & Henzel WJ (1989) *Biochem. Biophys. Res. Commun.* **161**, 851-858.
- Folkman J & Shing Y (1992) *J. Biol. Chem.* **267**, 10931-10934.
- Gospodarowicz D, Abraham JA & Schilling J (1989) *Proc. Natl. Acad. Sci USA* **86**, 7311-7315.
- Gospodarowicz D, Weseman J, Morgan JS & Lindstrom J (1976) *J. Cell Biol.*, **70**: 395-405.
- Hagg T, Quon D, Higaki J & Varon S (1992) *Neuron*, **8**, 145-158.
- Hefti S (1986) *J. Neurosci.*, **6**, 2155-2162.
- Hendry IA & Campbell J (1976) *J. Neurocytol.*, **5**, 351-360.
- Hendry IA, Murphy M, Hilton DJ, Nicola NA & Bartlett PF (1992) *J. Neurosci.* **12**, 3427-3434.
- Hunt *et al.*, (1967) *Am. J. Surgery*, **114**: 302-307.
- Koch AE, Harlow LA, Haines GK, Amento EP, Unemoti EN, Wong WL, Pope RM, Ferrara N, (1994) *J. Immunol.* **152**, 4149-4156.
- Kromer AF (1987) *Science*, **235**, 214-216.
- Larsson C, Weber G, Kvanta E, Lewis C, Janson M, Jones C, Glaser T, Evans G, Nordenskjold M, (1992) *Hum. Genet.* **89**, 187-193.
- OZUS Leung DW, Cachianes G, Kuang W-J, Goeddel DV & Ferrara N (1989) *Science* **246**:1306-1309.

- Lowe C, Cornish J, Callon K, Martin TJ & Reid IR (1991) *J. Bone Mineral Res.*, **6**, 1277-1283.
- Lowe C, Cornish J, Martin TJ & Reid IR (1991) *Calcif. Tissue Int.*, **49**, 394-397.
- Martinou JC, Martinou I & Kato AC (1992) *Neuron*, **8**, 737-744.
- Midy V & Plouet J (1994) *Biochem. Biophys. Res. Commun.*, **199**: 380-386.
- Miles AA & Miles EM (1952) *J. Physiol. (Lond)* **118**:228-257.
- Montesano R, Vassalli JD, Baird A, Guillemin R & Orci, L (1986) *Proc. Natl. Acad. Sci USA*, **83**, 7297-7301.
- Otto D., Frotscher M & Unsicker K (1989) *J. Neurosci. Res.*, **22**, 83-91.
- Pepper MS, Ferrara N, Orci L, Montesano R. (1991) *Biochem. Biophys. Res. Commun.* **181**, 902-906).
- Roth S & Weston J (1967) *Proc. Natl. Acad. Sci USA*, **58**: 974-980.
- Sambrook J, Fritsch EF, Maniatis T, (1989) *Molecular Cloning: A Laboratory Manual - 2nd Ed.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Schilling *et al.*, (1959) *Surgery*, **46**: 702-710.
- Senger DR, Van De Water L, Brown LF, Nagy JA, Yeo KT, Yeo TK, Berse B, Jackman RW, Dvorak AM, Dvorak HF (1993) *Cancer Metastasis Rev.* **12**, 303-324.
- Sharkey AM, Chamock-Jones DS, Boocock CA, Brown KD, Smith SK, (1993) *J. Reprod. Fertil.* **99**, 609-615.
- Sunderkotter C, Steinbrink K, Goebeler M, Bhardway R, Sorg E, (1993) *J. Leukocyt. Biol.* **55**, 410-422.
- Suter U, Angst C, Tien C-L, Drinkwater CC, Lindsay RM and Shooter EM (1992) *J. Neurosci.*, **12**, 306-318.
- Tischer E, Mitchell R, Hartman T, Silva M, Gospodarowicz D, Fiddes JC, & Abraham J (1991) *J. Biol. Chem.* **266**, 11947-11954.
- Williams LR, Varon S, Peterson GM, Wictorin K, Fischer W, Bjorklund A & Gage FH (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9231-9235.
- Xu et al, 1992.**
- Yan Z, Weich HA, Bernart W, Breckwoldt M, Neulen J, (1993) *J. Clin. Endocrinol. Metab.* **77**, 1723-1725.
- Yip NK, Rich KM, Lampe PA & Johnson EM Jr (1984) *J. Neurosci.*, **4**, 2986-2992.



## SEQUENCE LISTING

### (1) GENERAL INFORMATION:

(i) APPLICANT: THE COUNCIL OF THE QUEENSLAND INSTITUTE  
OF MEDICAL RESEARCH and AMRAD  
CORPORATION LIMITED

(ii) TITLE OF INVENTION: A NOVEL GROWTH FACTOR AND A  
GENETIC SEQUENCE ENCODING SAME

(iii) NUMBER OF SEQUENCES: 10

#### (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: DAVIES COLLISON CAVE

(B) STREET: 1 LITTLE COLLINS STREET

(C) CITY: MELBOURNE

(D) STATE: VICTORIA

(E) COUNTRY: AUSTRALIA

(F) ZIP: 3000

#### (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

#### (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: AU PROVISIONAL

(B) FILING DATE:

#### (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: HUGHES DR, E JOHN L

(C) REFERENCE/DOCKET NUMBER: EJH/EK

#### (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: +61 3 9254 2777

(B) TELEFAX: +61 3 9254 2770

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 649 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 17...589

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TCGGGCCTCC GAAACC ATG AAC TTT CTG CTG TCT TGG GTG CAT TGG AGC	49
Met Asn Phe Leu Leu Ser Trp Val His Trp Ser	
1 5 10	
CTT GCC TTG CTG CTC TAC CTC CAC CAT GCC AAG TGG TCC CAG GCT GCA	97
Leu Ala Leu Leu Leu Tyr Leu His His Ala Lys Trp Ser Gln Ala Ala	
15 20 25	
CCC ATG GCA GAA GGA GGA GGG CAG AAT CAT CAC GAA GTG GTG AAG TTC	145
Pro Met Ala Glu Gly Gly Gly Gln Asn His His Glu Val Val Lys Phe	
30 35 40	
ATG GAT GTC TAT CAG CGC AGC TAC TGC CAT CCA ATC GAG ACC CTG GTG	193
Met Asp Val Tyr Gln Arg Ser Tyr Cys His Pro Ile Glu Thr Leu Val	
45 50 55	
GAC ATC TTC CAG GAG TAC CCT GAT GAG ATC GAG TAC ATC TTC AAG CCA	241
Asp Ile Phe Gln Glu Tyr Pro Asp Glu Ile Glu Tyr Ile Phe Lys Pro	
60 65 70 75	
TCC TGT GTG CCC CTG ATG CGA TGC GGG GGC TGC TGC AAT GAC GAG GGC	289
Ser Cys Val Pro Leu Met Arg Cys Gly Gly Cys Cys Asn Asp Glu Gly	
80 85 90	
CTG GAG TGT GTG CCC ACT GAG GAG TCC AAC ATC ACC ATG CAG ATT ATG	337
Leu Glu Cys Val Pro Thr Glu Glu Ser Asn Ile Thr Met Gln Ile Met	
95 100 105	
CGG ATC AAA CCT CAC CAA GGC CAG CAC ATA GGA GAG ATG AGC TTC CTA	385
Arg Ile Lys Pro His Gln Gly Gln His Ile Gly Glu Met Ser Phe Leu	
110 115 120	
CAG CAC AAC AAA TGT GAA TGC AGA CCA AAG AAA GAT AGA GCA AGA CAA	433
Gln His Asn Lys Cys Glu Cys Arg Pro Lys Lys Asp Arg Ala Arg Gln	
125 130 135	

(i) SEQUENCE CHARACTERISTICS:

- (ii) MOLECULE TYPE: protein

## 951110,p:\oper\ejh\_ANGF.prv,23

Glu	Cys	Arg	Pro	Lys	Lys	Asp	Arg	Ala	Arg	Gln	Glu	Asn	Pro	Cys	Gly
130						135					140				
Pro	Cys	Ser	Glu	Arg	Arg	Lys	His	Leu	Phe	Val	Gln	Asp	Pro	Gln	Thr
145						150				155					160
Cys	Lys	Cys	Ser	Cys	Lys	Asn	Thr	Asp	Ser	Arg	Cys	Lys	Ala	Arg	Gln
				165					170					175	
Leu	Glu	Leu	Asn	Glu	Arg	Thr	Cys	Arg	Cys	Asp	Lys	Pro	Arg	Arg	
			180					185					190		

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1094 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..624

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CC ATG AGC CCT CTG CTC CGC CGC CTG CTG CTC GCC GCA CTC CTG CAG	47
Met Ser Pro Leu Leu Arg Arg Leu Leu Leu Ala Ala Leu Leu Gln	
1 5 10 15	
CTG GCC CCC GCC CAG GCC CCT GTC TCC CAG CCT GAT GCC CCT GGC CAC	95
Leu Ala Pro Ala Gln Ala Pro Val Ser Gln Pro Asp Ala Pro Gly His	
20 25 30	
CAG AGG AAA GTG GTG TCA TGG ATA GAT GTG TAT ACT CGC GCT ACC TGC	143
Gln Arg Lys Val Val Ser Trp Ile Asp Val Tyr Thr Arg Ala Thr Cys	
35 40 45	
CAG CCC CGG GAG GTG GTG GTG CCC TTG ACT GTG GAG CTC ATG GGC ACC	191
Gln Pro Arg Glu Val Val Val Pro Leu Thr Val Glu Leu Met Gly Thr	
50 55 60	
GTG GCC AAA CAG CTG GTG CCC AGC TGC GTG ACT GTG CAG CGC TGT GGT	239
Val Ala Lys Gln Leu Val Pro Ser Cys Val Thr Val Gln Arg Cys Gly	
65 70 75	
GGC TGC TGC CCT GAC GAT GGC CTG GAG TGT GTG CCC ACT GGG CAG CAC	287
Gly Cys Cys Pro Asp Asp Gly Leu Glu Cys Val Pro Thr Gly Gln His	
80 85 90 95	

CAA GTC CGG ATG CAG ATC CTC ATG ATC CGG TAC CCG AGC AGT CAG CTG	335
Gln Val Arg Met Gln Ile Leu Met Ile Arg Tyr Pro Ser Ser Gln Leu	
100 105 110	
GGG GAG ATG TCC CTG GAA GAA CAC AGC CAG TGT GAA TGC AGA CCT AAA	383
Gly Glu Met Ser Leu Glu Glu His Ser Gln Cys Glu Cys Arg Pro Lys	
115 120 125	
AAA AAG GAC AGT GCT GTG AAG CCA GAC AGG GCT GCC ACT CCC CAC CAC	431
Lys Lys Asp Ser Ala Val Lys Pro Asp Arg Ala Ala Thr Pro His His	
130 135 140	
CGT CCC CAG CCC CGT TCT GTT CCG GGC TGG GAC TCT GCC CCC GGA GCA	479
Arg Pro Gln Pro Arg Ser Val Pro Gly Trp Asp Ser Ala Pro Gly Ala	
145 150 155	
CCC TCC CCA GCT GAC ATC ACC CAT CCC ACT CCA GCC CCA GGC CCC TCT	527
Pro Ser Pro Ala Asp Ile Thr His Pro Thr Pro Ala Pro Gly Pro Ser	
160 165 170 175	
GCC CAC GCT GCA CCC AGC ACC ACC AGC GCC CTG ACC CCC GGA CCT GCC	575
Ala His Ala Ala Pro Ser Thr Thr Ser Ala Leu Thr Pro Gly Pro Ala	
180 185 190	
GCT GCC GCT GCC GAC GCC GCA GCT TCC TCC GTT GCC AAG GGC GGG GCT T	624
Ala Ala Ala Ala Asp Ala Ala Ala Ser Ser Val Ala Lys Gly Gly Ala	
195 200 205	
AGAGCTCAAC CCAGACACCT GCAGGTGCCG GAAGCTGCCA AGGTGACACA TGGCTTTTCA	684
GACTCAGCAG GGTGACTTGC CTCAGAGGCT ATATCCCAGT GGGGGAACAA AGGGGAGCCT	744
GGTAAAAAAC AGCCAAGCCC CCAAGACCTC AGCCCAGGCA GAAGCTGCTC TAGGACCTGG	804
GCCTCTCAGA GGGCTCTTCT GCCATCCCTT GTCTCCCTGA GGCCATCATC AAACAGGACA	864
GAGTTGGAAG AGGAGACTGG GAGGCAGCAA GAGGGGTCAC ATACCAGCTC AGGGGAGAAT	924
GGAGTACTGT CTCAGTTTCT AACCACTCTG TGCAAGTAAG CATCTTACAA CTGGCTCTTC	984
CTCCCCCTCAC TAAGAAGACC CAAACCTCTG CATAATGGGA TTTGGGCTTT GGTACAAGAA	1044
CTGTGACCCC CAACCCTGAT AAAAGAGATG GAAGGAAAAA AAAAAAAAAA	1094

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 207 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Ser Pro Leu Leu Arg Arg Leu Leu Leu Ala Ala Leu Leu Gln Leu
 1              5              10              15

Ala Pro Ala Gln Ala Pro Val Ser Gln Pro Asp Ala Pro Gly His Gln
      20              25              30

Arg Lys Val Val Ser Trp Ile Asp Val Tyr Thr Arg Ala Thr Cys Gln
      35              40              45

Pro Arg Glu Val Val Val Pro Leu Thr Val Glu Leu Met Gly Thr Val
      50              55              60

Ala Lys Gln Leu Val Pro Ser Cys Val Thr Val Gln Arg Cys Gly Gly
      65              70              75              80

Cys Cys Pro Asp Asp Gly Leu Glu Cys Val Pro Thr Gly Gln His Gln
      85              90              95

Val Arg Met Gln Ile Leu Met Ile Arg Tyr Pro Ser Ser Gln Leu Gly
      100              105              110

Glu Met Ser Leu Glu Glu His Ser Gln Cys Glu Cys Arg Pro Lys Lys
      115              120              125

Lys Asp Ser Ala Val Lys Pro Asp Arg Ala Ala Thr Pro His His Arg
      130              135              140

Pro Gln Pro Arg Ser Val Pro Gly Trp Asp Ser Ala Pro Gly Ala Pro
      145              150              155              160

Ser Pro Ala Asp Ile Thr His Pro Thr Pro Ala Pro Gly Pro Ser Ala
      165              170              175

His Ala Ala Pro Ser Thr Thr Ser Ala Leu Thr Pro Gly Pro Ala Ala
      180              185              190

Ala Ala Ala Asp Ala Ala Ala Ser Ser Val Ala Lys Gly Gly Ala
      195              200              205

```

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 993 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 3..566

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CC ATG AGC CCT CTG CTC CGC CGC CTG CTG CTC GCC GCA CTC CTG CAG	47
Met Ser Pro Leu Leu Arg Arg Leu Leu Leu Ala Ala Leu Leu Gln	
1 5 10 15	
CTG GCC CCC GCC CAG GCC CCT GTC TCC CAG CCT GAT GCC CCT GGC CAC	95
Leu Ala Pro Ala Gln Ala Pro Val Ser Gln Pro Asp Ala Pro Gly His	
20 25 30	
CAG AGG AAA GTG GTG TCA TGG ATA GAT GTG TAT ACT CGC GCT ACC TGC	143
Gln Arg Lys Val Val Ser Trp Ile Asp Val Tyr Thr Arg Ala Thr Cys	
35 40 45	
CAG CCC CGG GAG GTG GTG GTG CCC TTG ACT GTG GAG CTC ATG GGC ACC	191
Gln Pro Arg Glu Val Val Val Pro Leu Thr Val Glu Leu Met Gly Thr	
50 55 60	
GTG GCC AAA CAG CTG GTG CCC AGC TGC GTG ACT GTG CAG CGC TGT GGT	239
Val Ala Lys Gln Leu Val Pro Ser Cys Val Thr Val Gln Arg Cys Gly	
65 70 75	
GGC TGC TGC CCT GAC GAT GGC CTG GAG TGT GTG CCC ACT GGG CAG CAC	287
Gly Cys Cys Pro Asp Asp Gly Leu Glu Cys Val Pro Thr Gly Gln His	
80 85 90 95	
CAA GTC CGG ATG CAG ATC CTC ATG ATC CGG TAC CCG AGC AGT CAG CTG	335
Gln Val Arg Met Gln Ile Leu Met Ile Arg Tyr Pro Ser Ser Gln Leu	
100 105 110	
GGG GAG ATG TCC CTG GAA GAA CAC AGC CAG TGT GAA TGC AGA CCT AAA	383
Gly Glu Met Ser Leu Glu Glu His Ser Gln Cys Glu Cys Arg Pro Lys	
115 120 125	
AAA AAG GAC AGT GCT GTG AAG CCA GAT AGC CCC AGG CCC CTC TGC CCA	431
Lys Lys Asp Ser Ala Val Lys Pro Asp Ser Pro Arg Pro Leu Cys Pro	
130 135 140	

CGC TGC ACC CAG CAC CAC CAG CGC CCT GAC CCC CGG ACC TGC CGC TGC	479
Arg Cys Thr Gln His His Gln Arg Pro Asp Pro Arg Thr Cys Arg Cys	
145 150 155	
CGC TGC CGA CGC CGC AGC TTC CTC CGT TGC CAA GGG CGG GGC TTA GAG	527
Arg Cys Arg Arg Arg Ser Phe Leu Arg Cys Gln Gly Arg Gly Leu Glu	
160 165 170 175	
CTC AAC CCA GAC ACC TGC AGG TGC CGG AAG CTG CGA AGG TGACACATGG	576
Leu Asn Pro Asp Thr Cys Arg Cys Arg Lys Leu Arg Arg	
180 185	
CTTTTCAGAC TCAGCAGGGT GACTTGCCTC AGAGGCTATA TCCCAGTGGG GGAACAAAGG	636
GGAGCCTGGT AAAAAACAGC CAAGCCCCCA AGACCTCAGC CCAGGCAGAA GCTGCTCTAG	696
GACCTGGGCC TCTCAGAGGG CTCTTCTGCC ATCCCTTGTC TCCCTGAGGC CATCATCAAA	756
CAGGACAGAG TTGGAAGAGG AGACTGGGAG GCAGCAAGAG GGGTCACATA CCAGCTCAGG	816
GGAGAATGGA GTACTGTCTC AGTTTCTAAC CACTCTGTGC AAGTAAGCAT CTTACAACTG	876
GCTCTTCCTC CCCTCACTAA GAAGACCCAA ACCTCTGCAT AATGGGATTT GGGCTTTGGT	936
ACAAGAACTG TGACCCCCAA CCCTGATAAA AGAGATGGAA GGAAAAAAAA AAAAAAA	993

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 188 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Ser Pro Leu Leu Arg Arg Leu Leu Leu Ala Ala Leu Leu Gln Leu	
1 5 10 15	
Ala Pro Ala Gln Ala Pro Val Ser Gln Pro Asp Ala Pro Gly His Gln	
20 25 30	
Arg Lys Val Val Ser Trp Ile Asp Val Tyr Thr Arg Ala Thr Cys Gln	
35 40 45	
Pro Arg Glu Val Val Val Pro Leu Thr Val Glu Leu Met Gly Thr Val	
50 55 60	
Ala Lys Gln Leu Val Pro Ser Cys Val Thr Val Gln Arg Cys Gly Gly	
65 70 75 80	



Cys	Cys	Pro	Asp	Asp	Gly	Leu	Glu	Cys	Val	Pro	Thr	Gly	Gln	His	Gln	
				85					90					95		
Val	Arg	Met	Gln	Ile	Leu	Met	Ile	Arg	Tyr	Pro	Ser	Ser	Gln	Leu	Gly	
			100					105					110			
Glu	Met	Ser	Leu	Glu	Glu	His	Ser	Gln	Cys	Glu	Cys	Arg	Pro	Lys	Lys	
		115					120					125				
Lys	Asp	Ser	Ala	Val	Lys	Pro	Asp	Ser	Pro	Arg	Pro	Leu	Cys	Pro	Arg	
	130					135					140					
Cys	Thr	Gln	His	His	Gln	Arg	Pro	Asp	Pro	Arg	Thr	Cys	Arg	Cys	Arg	
145					150					155					160	
Cys	Arg	Arg	Arg	Ser	Phe	Leu	Arg	Cys	Gln	Gly	Arg	Gly	Leu	Glu	Leu	
				165					170					175		
Asn	Pro	Asp	Thr	Cys	Arg	Cys	Arg	Lys	Leu	Arg	Arg					
			180					185								

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 858 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 3..431

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CC ATG AGC CCT CTG CTC CGC CGC CTG CTG CTC GCC GCA CTC CTG CAG	47
Met Ser Pro Leu Leu Arg Arg Leu Leu Leu Ala Ala Leu Leu Gln	
1 5 10 15	
CTG GCC CCC GCC CAG GCC CCT GTC TCC CAG CCT GAT GCC CCT GGC CAC	95
Leu Ala Pro Ala Gln Ala Pro Val Ser Gln Pro Asp Ala Pro Gly His	
20 25 30	
CAG AGG AAA GTG GTG TCA TGG ATA GAT GTG TAT ACT CGC GCT ACC TGC	143
Gln Arg Lys Val Val Ser Trp Ile Asp Val Tyr Thr Arg Ala Thr Cys	
35 40 45	
CAG CCC CGG GAG GTG GTG GTG CCC TTG ACT GTG GAG CTC ATG GGC ACC	191
Gln Pro Arg Glu Val Val Val Pro Leu Thr Val Glu Leu Met Gly Thr	
50 55 60	

GTG GCC AAA CAG CTG GTG CCC AGC TGC GTG ACT GTG CAG CGC TGT GGT	239
Val Ala Lys Gln Leu Val Pro Ser Cys Val Thr Val Gln Arg Cys Gly	
65 70 75	
GGC TGC TGC CCT GAC GAT GGC CTG GAG TGT GTG CCC ACT GGG CAG CAC	287
Gly Cys Cys Pro Asp Asp Gly Leu Glu Cys Val Pro Thr Gly Gln His	
80 85 90 95	
CAA GTC CGG ATG CAG ATC CTC ATG ATC CGG TAC CCG AGC AGT CAG CTG	335
Gln Val Arg Met Gln Ile Leu Met Ile Arg Tyr Pro Ser Ser Gln Leu	
100 105 110	
GGG GAG ATG TCC CTG GAA GAA CAC AGC CAG TGT GAA TGC AGA CCT AAA	383
Gly Glu Met Ser Leu Glu Glu His Ser Gln Cys Glu Cys Arg Pro Lys	
115 120 125	
AAA AAG GAC AGT GCT GTG AAG CCA GAT AGG TGC CGG AAG CTG CGA AGG	431
Lys Lys Asp Ser Ala Val Lys Pro Asp Arg Cys Arg Lys Leu Arg Arg	
130 135 140	
TGACACATGG CTTTTTCAGAC TCAGCAGGGT GACTTGCCTC AGAGGCTATA TCCCAGTGGG	491
GGAACAAAGG GGAGCCTGGT AAAAAACAGC CAAGCCCCCA AGACCTCAGC CCAGGCAGAA	551
GCTGCTCTAG GACCTGGGCC TCTCAGAGGG CTCTTCTGCC ATCCCTTGTC TCCCTGAGGC	611
CATCATCAAA CAGGACAGAG TTGGAAGAGG AGACTGGGAG GCAGCAAGAG GGGTCACATA	671
CCAGCTCAGG GGAGAATGGA GTACTGTCTC AGTTTCTAAC CACTCTGTGC AAGTAAGCAT	731
CTTACAACCTG GCTCTTCCTC CCCTCACTAA GAAGACCCAA ACCTCTGCAT AATGGGATTT	791
GGGCTTTGGT ACAAGAACTG TGACCCCCAA CCCTGATAAA AGAGATGGAA GGAAAAAAAA	851
AAAAAAA	858

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 143 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met	Ser	Pro	Leu	Leu	Arg	Arg	Leu	Leu	Leu	Ala	Ala	Leu	Leu	Gln	Leu	
1				5					10					15		
Ala	Pro	Ala	Gln	Ala	Pro	Val	Ser	Gln	Pro	Asp	Ala	Pro	Gly	His	Gln	
			20					25					30			
Arg	Lys	Val	Val	Ser	Trp	Ile	Asp	Val	Tyr	Thr	Arg	Ala	Thr	Cys	Gln	
		35					40					45				
Pro	Arg	Glu	Val	Val	Val	Pro	Leu	Thr	Val	Glu	Leu	Met	Gly	Thr	Val	
	50					55					60					
Ala	Lys	Gln	Leu	Val	Pro	Ser	Cys	Val	Thr	Val	Gln	Arg	Cys	Gly	Gly	
65					70					75					80	
Cys	Cys	Pro	Asp	Asp	Gly	Leu	Glu	Cys	Val	Pro	Thr	Gly	Gln	His	Gln	
				85					90					95		
Val	Arg	Met	Gln	Ile	Leu	Met	Ile	Arg	Tyr	Pro	Ser	Ser	Gln	Leu	Gly	
			100					105					110			
Glu	Met	Ser	Leu	Glu	Glu	His	Ser	Gln	Cys	Glu	Cys	Arg	Pro	Lys	Lys	
		115					120					125				
Lys	Asp	Ser	Ala	Val	Lys	Pro	Asp	Arg	Cys	Arg	Lys	Leu	Arg	Arg		
	130					135					140					

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 910 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 3..305

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CC ATG AGC CCT CTG CTC CGC CGC CTG CTG CTC GCC GCA CTC CTG CAG	47
Met Ser Pro Leu Leu Arg Arg Leu Leu Ala Ala Leu Leu Gln	
1 5 10 15	
CTG GCC CCC GCC CAG GCC CCT GTC TCC CAG CCT GAT GCC CCT GGC CAC	95
Leu Ala Pro Ala Gln Ala Pro Val Ser Gln Pro Asp Ala Pro Gly His	
20 25 30	
CAG AGG AAA GTG GTG TCA TGG ATA GAT GTG TAT ACT CGC GCT ACC TGC	143
Gln Arg Lys Val Val Ser Trp Ile Asp Val Tyr Thr Arg Ala Thr Cys	
35 40 45	
CAG CCC CGG GAG GTG GTG GTG CCC TTG ACT GTG CAG CTC ATG GGC ACC	191
Gln Pro Arg Glu Val Val Val Pro Leu Thr Val Glu Leu Met Gly Thr	
50 55 60	
GTG GCC AAA CAG CTG GTG CCC AGC TGC GTG ACT GTG CAG CGC TGT GGT	239
Val Ala Lys Gln Leu Val Pro Ser Cys Val Thr Val Gln Arg Cys Gly	
65 70 75	
GGC TGC TGC CCT GAC GAT GGC CTG GAG TGT GTG CCC ACT GGG CAG CAC	287
Gly Cys Cys Pro Asp Asp Gly Leu Glu Cys Val Pro Thr Gly Gln His	
80 85 90 95	
CAA GTC CGG ATG CAG ACC TAAAAAAAAG GACAGTGCTG TGAAGCCAGA	335
Gln Val Arg Met Gln Thr	
100	
CAGGGCTGCC ACTCCCCACC ACCGTCCCCA GCCCCGTTCT GTTCCGGGCT GGGACTCTGC	395
CCCCGGAGCA CCCTCCCCAG CTGACATCAC CCATCCCACT CCAGCCCCAG GCCCCTCTGC	455
CCACGCTGCA CCCAGCACCA CCAGCGCCCT GACCCCCGGA CCTGCCGCTG CCGCTGCCGA	515
CGCCGCAGCT TCCTCCGTTG CCAAGGGCGG GGCTTAGAGC TCAACCCAGA CACCTGCAGG	575
TGCCGGAAGC TGCGAAGGTG ACACATGGCT TTTCAGACTC AGCAGGGTGA CTTGCCTCAG	635

AGGCTATATC CCAGTGGGGA ACAAAGAGGA GCCTGGTAAA AAACAGCCAA GCCCCAAGA 695  
CCTCAGCCCA GGCAGAAGCT GCTCTAGGAC CTGGGCCTCT CAGAGGGCTC TTCTGCCATC 755  
CCTTGTCTCC CTGAGGCCAT CATCAAACAG GACAGAGTTG GAAGAGGAGA CTGGGAGGCA 815  
GCAAGAGGGG TCACATACCA GCTCAGGGGA GAATGGAGTA CTGTCTCAGT TTCTAACCAC 875  
TCTGTGCAAG TAAGCATCTT ACAACTGGCT CTTCC 910

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 101 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Ser Pro Leu Leu Arg Arg Leu Leu Leu Ala Ala Leu Leu Gln Leu  
1 5 10 15  
Ala Pro Ala Gln Ala Pro Val Ser Gln Pro Asp Ala Pro Gly His Gln  
20 25 30  
Arg Lys Val Val Ser Trp Ile Asp Val Tyr Thr Arg Ala Thr Cys Gln  
35 40 45  
Pro Arg Glu Val Val Val Pro Leu Thr Val Glu Leu Met Gly Thr Val  
50 55 60  
Ala Lys Gln Leu Val Pro Ser Cys Val Thr Val Gln Arg Cys Gly Gly  
65 70 75 80  
Cys Cys Pro Asp Asp Gly Leu Glu Cys Val Pro Thr Gly Gln His Gln  
85 90 95  
Val Arg Met Gln Thr  
100

DATED this 20th day of November, 1995

THE COUNCIL OF THE QUEENSLAND INSTITUTE OF MEDICAL RESEARCH  
and AMRAD CORPORATION LIMITED

By Its Patent Attorneys

DAVIES COLLISON CAVE

FIGURE 1

TCGGGCCTCC GAAACC ATG AAC TTT CTG CTG TCT TGG GTG CAT TGG AGC	49
Met Asn Phe Leu Leu Ser Trp Val His Trp Ser	
1 5 10	
CTT GCC TTG CTG CTC TAC CTC CAC CAT GCC AAG TGG TCC CAG GCT GCA	97
Leu Ala Leu Leu Leu Tyr Leu His His Ala Lys Trp Ser Gln Ala Ala	
15 20 25	
CCC ATG GCA GAA GGA GGA GGG CAG AAT CAT CAC GAA GTG GTG AAG TTC	145
Pro Met Ala Glu Gly Gly Gly Gln Asn His His Glu Val Val Lys Phe	
30 35 40	
ATG GAT GTC TAT CAG CGC AGC TAC TGC CAT CCA ATC GAG ACC CTG GTG	193
Met Asp Val Tyr Gln Arg Ser Tyr Cys His Pro Ile Glu Thr Leu Val	
45 50 55	
GAC ATC TTC CAG GAG TAC CCT GAT GAG ATC GAG TAC ATC TTC AAG CCA	241
Asp Ile Phe Gln Glu Tyr Pro Asp Glu Ile Glu Tyr Ile Phe Lys Pro	
60 65 70 75	
TCC TGT GTG CCC CTG ATG CGA TGC GGG GGC TGC TGC AAT GAC GAG GGC	289
Ser Cys Val Pro Leu Met Arg Cys Gly Gly Cys Cys Asn Asp Glu Gly	
80 85 90	
CTG GAG TGT GTG CCC ACT GAG GAG TCC AAC ATC ACC ATG CAG ATT ATG	337
Leu Glu Cys Val Pro Thr Glu Glu Ser Asn Ile Thr Met Gln Ile Met	
95 100 105	
CGG ATC AAA CCT CAC CAA GGC CAG CAC ATA GGA GAG ATG AGC TTC CTA	385
Arg Ile Lys Pro His Gln Gly Gln His Ile Gly Glu Met Ser Phe Leu	
110 115 120	
CAG CAC AAC AAA TGT GAA TGC AGA CCA AAG AAA GAT AGA GCA AGA CAA	433
Gln His Asn Lys Cys Glu Cys Arg Pro Lys Lys Asp Arg Ala Arg Gln	
125 130 135	
GAA AAT CCC TGT GGG CCT TGC TCA GAG CGG AGA AAG CAT TTG TTT GTA	481
Glu Asn Pro Cys Gly Pro Cys Ser Glu Arg Arg Lys His Leu Phe Val	
140 145 150 155	
CAA GAT CCG CAG ACG TGT AAA TGT TCC TGC AAA AAC ACA GAC TCG CGT	529
Gln Asp Pro Gln Thr Cys Lys Cys Ser Cys Lys Asn Thr Asp Ser Arg	
160 165 170	
TGC AAG GCG AGG CAG CTT GAG TTA AAC GAA CGT ACT TGC AGA TGT GAC	577
Cys Lys Ala Arg Gln Leu Glu Leu Asn Glu Arg Thr Cys Arg Cys Asp	
175 180 185	
AAG CCG AGG CGG TGAGCCGGGC AGGAGGAAGG AGCCTCCCTC AGCGTTTCGG	629
Lys Pro Arg Arg	
190	
GAACCAGATC TCTCACCAGG	649

FIGURE 2

CC ATG AGC CCT CTG CTC CGC CGC CTG CTG CTC GCC GCA CTC CTG CAG	47
Met Ser Pro Leu Leu Arg Arg Leu Leu Leu Ala Ala Leu Leu Gln	
1 5 10 15	
CTG GCC CCC GCC CAG GCC CCT GTC TCC CAG CCT GAT GCC CCT GGC CAC	95
Leu Ala Pro Ala Gln Ala Pro Val Ser Gln Pro Asp Ala Pro Gly His	
20 25 30	
CAG AGG AAA GTG GTG TCA TGG ATA GAT GTG TAT ACT CGC GCT ACC TGC	143
Gln Arg Lys Val Val Ser Trp Ile Asp Val Tyr Thr Arg Ala Thr Cys	
35 40 45	
CAG CCC CGG GAG GTG GTG GTG CCC TTG ACT GTG GAG CTC ATG GGC ACC	191
Gln Pro Arg Glu Val Val Val Pro Leu Thr Val Glu Leu Met Gly Thr	
50 55 60	
GTG GCC AAA CAG CTG GTG CCC AGC TGC GTG ACT GTG CAG CGC TGT GGT	239
Val Ala Lys Gln Leu Val Pro Ser Cys Val Thr Val Gln Arg Cys Gly	
65 70 75	
GGC TGC TGC CCT GAC GAT GGC CTG GAG TGT GTG CCC ACT GGG CAG CAC	287
Gly Cys Cys Pro Asp Asp Gly Leu Glu Cys Val Pro Thr Gly Gln His	
80 85 90 95	
CAA GTC CGG ATG CAG ATC CTC ATG ATC CGG TAC CCG AGC AGT CAG CTG	335
Gln Val Arg Met Gln Ile Leu Met Ile Arg Tyr Pro Ser Ser Gln Leu	
100 105 110	
GGG GAG ATG TCC CTG GAA GAA CAC AGC CAG TGT GAA TGC AGA CCT AAA	383
Gly Glu Met Ser Leu Glu Glu His Ser Gln Cys Glu Cys Arg Pro Lys	
115 120 125	
AAA AAG GAC AGT GCT GTG AAG CCA GAC AGG GCT GCC ACT CCC CAC CAC	431
Lys Lys Asp Ser Ala Val Lys Pro Asp Arg Ala Ala Thr Pro His His	
130 135 140	
CGT CCC CAG CCC CGT TCT GTT CCG GGC TGG GAC TCT GCC CCC GGA GCA	479
Arg Pro Gln Pro Arg Ser Val Pro Gly Trp Asp Ser Ala Pro Gly Ala	
145 150 155	
CCC TCC CCA GCT GAC ATC ACC CAT CCC ACT CCA GCC CCA GGC CCC TCT	527
Pro Ser Pro Ala Asp Ile Thr His Pro Thr Pro Ala Pro Gly Pro Ser	
160 165 170 175	
GCC CAC GCT GCA CCC AGC ACC ACC AGC GCC CTG ACC CCC GGA CCT GCC	575
Ala His Ala Ala Pro Ser Thr Thr Ser Ala Leu Thr Pro Gly Pro Ala	
180 185 190	
GCT GCC GCT GCC GAC GCC GCA GCT TCC TCC GTT GCC AAG GGC GGG GCT T	624
Ala Ala Ala Ala Asp Ala Ala Ala Ser Ser Val Ala Lys Gly Gly Ala	
195 200 205	

FIGURE 2 (continued...)

AGAGCTCAAC CCAGACACCT GCAGGTGCCG GAAGCTGCCA AGGTGACACA TGGCTTTTCA	684
GA CTCAGCAG GGTGACTTGC CTCAGAGGCT ATATCCCAGT GGGGGAACAA AGGGGAGCCT	744
GGTAAAAAAC AGCCAAGCCC CCAAGACCTC AGCCCAGGCA GAAGCTGCTC TAGGACCTGG	804
GCCTCTCAGA GGGCTCTTCT GCCATCCCTT GTCTCCCTGA GGCCATCATC AAACAGGACA	864
GAGTTGGAAG AGGAGACTGG GAGGCAGCAA GAGGGGTCAC ATACCAGCTC AGGGGAGAAT	924
GGAGTACTGT CTCAGTTTCT AACCCTCTG TGCAAGTAAG CATCTTACAA CTGGCTCTTC	984
CTCCCCTCAC TAAGAAGACC CAAACCTCTG CATAATGGGA TTTGGGCTTT GGTACAAGAA	1044
CTGTGACCCC CAACCCTGAT AAAAGAGATG GAAGGAAAAA AAAAAAAAAA	1094



### FIGURE 3

>VEGF\_HUMAN VEGF\_HUMAN VASCULAR ENDOTHELIAL GROWTH FACTOR PRECURSOR (VEGF)  
(VASCULAR 215 AA.

Length = 215

Score = 181 (92.4 bits), Expect = 6.4e-20, P = 6.4e-20

Identities = 33/75 (44%), Positives = 48/75 (64%)

Query: 31 HQRKVSWIDVYTRATCQPREVVVPLTVELMGTVAQLVPSCVTVQRCGGCCPDDGLECV 90

+++ VV +DVY R+ C+P E +V + E + PSCV + RCGGCC D+GLECV

Sbjct: 36 NHHEVVKFMDVYQRSYCHPIETLVDIFQEYPDEIEYIFKPSCVPLMRCGGCCNDEGLECV 95

Query: 91 PTGQHQVRMQILMIR 105

PT + + MQI+ I+

Sbjct: 96 PTEESNITMQIMRIK 110

Score = 76 (38.8 bits), Expect = 0.0011, Poisson P(2) = 9.1e-12

Identities = 12/19 (63%), Positives = 16/19 (84%)

Query: 110 QLGEMSLEEHSQCECRPKK 128

++GEMS +H+ CECRPKK

Sbjct: 116 HIGEMSFLQHNKCECRPKK 134

Score = 72 (36.8 bits), Expect = 0.0046, Poisson P(3) = 3.6e-18

Identities = 14/21 (66%), Positives = 15/21 (71%)

Query: 202 RCQGRGLELNPDTCRCRKLRR 222

RC +R LELN TCRC K RR

Sbjct: 195 RCKARQLELNERTCRC DKPRR 215

Score = 46 (23.5 bits), Expect = 47., Poisson P(4) = 7.3e-10

Identities = 6/10 (60%), Positives = 9/10 (90%)

Query: 187 DPRTCRCRCR 196

DP+TC+C C+

Sbjct: 181 DPQTCCKCSCK 190

# FIGURE 4

Gap Weight: 3.000 Average Match: 1.000  
 Length Weight: 0.100 Average Mismatch: -0.900  
 Quality: 100.9 Length: 739  
 Ratio: 0.175 Gaps: 30  
 Percent Similarity: 69.703 Percent Identity: 69.703

```

28 ATGAGCCCTCTGCTCCGCGCCTGCTGCTCGCCGCACT ... 67
   ||||| | ||||| | || | | |||||
17 ATGAACCTTCTGCT....GTCT....TGGGTGCATTGGAGCCTTGCC 56

68 TGCAGCTGGCCCCCGCCAGGCCCTGTCTCCAGCCTGATGCCCTTGGC 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
57 TGCTGCTCTACCTCCACCATGCCAAGTGGTCCAGGCTGCA.CCCATGGC 105

118 CACCAGAGGA.....AAGTGGTG....TCATGGATAGAT 147
     ||||| ||||| ||||| ||||| |||
106 AGAAGGAGGAGGGCAGAATCATCACGAAGTGGTGAAGTTCATG....GAT 151

148 GTGTATACTCGC.GCTACCTGCCAGCCCCCGGAG...GTGGTGGTGGCCT 193
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
152 GTCTATCAGCGCAGCTA.CTGCCATCCAATCGAGACCTGGTGGACATCT 200

194 T....GA....CTGTGGAGGTCATGGGCACCGTGGCCAAACAGCTGGTG 234
   | ||| ||| ||| ||| ||| ||| ||| ||| |||
201 TCCAGGAGTACCCTGATGAGATCGAGTACATCTT...CAA.... 238

235 CCCAGCTGCCGTGACTGTGCAGCGCTG*GGTGGCTGCTGCCCTGACGATGG 284
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
239 CCATCCTGTGTGCCCTGATGCGATGCGGGGGCTGCTGCAATGACGAGGG 288

285 CCTGGAGTGTGTGCCCACTGGGAGCAGCAAGTCCGGATGCAGAT ... 329
   ||| ||||| ||||| ||||| ||| ||| ||| ||| |||
289 CCTGGAGTGTGTGCCCACTGAGGAGTCCAACATCACCATGCAGATTATGC 338

330 .....CCTCATGATCCGGTACCCGAGCAGTCAGC...TGGGGGAGAT 368
   ||||| ||||| ||||| ||||| ||| ||| ||| |||
339 GGATCAAAACCTCA...CCAAAG...GCCAGCACATAGGAGAGAT 375

369 GTCCCTGGAAAGAACACAGCCAGTGTGAATGCAGACCTAAAAAAAAGGACA 418
   | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
376 GAGCTTCCTACAGCACAACAAATGTGAATGCAGACC...AAAGAAAGATA 422

419 GTGCTGTGAAGCCAGACAGGGCTGCCACTCCCCAGCACCCTCCCCAGCCC 468
   ||| ||||| ||||| ||||| ||| ||| ||| ||| |||
423 G.....AGCAAGACAAG...AAAAATCCC... 442

469 CGTTCTGTTCGGGGCTGGGACTCTGCCCGCGGAGCACCTCCCCAGCTGA 518
   || ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
443 .TGTGGGCTTGC*GAGA...GGGAGAA 467

519 CATCAGCCATCCCACTCCAGCCCCAGGGCCCTCTGCCACGCTGCACCCA 568

468 ..... 468

569 GC.....ACCACAGCGCCCTGACCGCGGAGCCTGCCGCTGCCGC 608
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
469 GCATTTGTTTGTACAA...GATCCGCAGAGTGTAAATGTTCC 508

609 TGGCGAGCGCCGAGCTTCTTCCGTTGCCAAGGGCGGGGC...TAGAGCTC 656
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
509 TG.CAAAAACACAGACTC...CGTTGC...AAGGCGAGGAGCTTGAATT 553

657 AACCCAGAGACCTTGCAGGTGCCCGAAGCTGCCAAGGTTGA 695

554 AACGAAGCTATTGCAAGTGTGACAAAGCGAGAGCGAGTGA 692
  
```

FIGURE 5

165SOMS.Q.MSF.msf MSF: 687 Type: D Tuesday, June 20, 1995 Check: 3140

```

1
VEGF165 ATGAACTTTCTGCTGCTTGGGTGCATTGGAGCCTTGCCCTTGCTGCTCTACCTCCACCATGCCAAGTGGTCCCAGGCTG.
SOM175 ATGAGCCCTCTGCTCCGCCGCCTGCTGCTCGCCGCACTCCTGCAGCTGGCCCCCGCCAGGCCCCCTGTCTCCCAGCCTGA
SOM175-e6 ATGAGCCCTCTGCTCCGCCGCCTGCTGCTCGCCGCACTCCTGCAGCTGGCCCCCGCCAGGCCCCCTGTCTCCCAGCCTGA
SOM175-e6&7 ATGAGCCCTCTGCTCCGCCGCCTGCTGCTCGCCGCACTCCTGCAGCTGGCCCCCGCCAGGCCCCCTGTCTCCCAGCCTGA
SOM175-e4 ATGAGCCCTCTGCTCCGCCGCCTGCTGCTCGCCGCACTCCTGCAGCTGGCCCCCGCCAGGCCCCCTGTCTCCCAGCCTGA

81
VEGF165 CACCCATGGCAGAAGGAGGAGGGCAGAATCATCAGAAAGTGGTGAAGTTCATGGATGTCTATCAGCGCAGCTACTGCCAT
SOM175 TGCCCCCTGGCCACCAGAGGAAAGTGGTGTATGGATAGATGTGTATACTCGCG....CTACCTGC.CAGCC.CCGGGAG
SOM175-e6 TGCCCCCTGGCCACCAGAGGAAAGTGGTGTATGGATAGATGTGTATACTCGCG....CTACCTGC.CAGCC.CCGGGAG
SOM175-e6&7 TGCCCCCTGGCCACCAGAGGAAAGTGGTGTATGGATAGATGTGTATACTCGCG....CTACCTGC.CAGCC.CCGGGAG
SOM175-e4 TGCCCCCTGGCCACCAGAGGAAAGTGGTGTATGGATAGATGTGTATACTCGCG....CTACCTGC.CAGCC.CCGGGAG

161
VEGF165 CCAATCGAGACCCTGGTGGACATCTTCCAGGAGTACCCTGATGAGATCGAGTACATCTTCAAGCCATCCTGTGTGCCCT
SOM175 GTGGTGGTGGCCCTTGACTG.TGGAGCTCATGGGCACCGTGGCCAAAC..AGCTGGTGCCAG.....CTGCGTGACTGT
SOM175-e6 GTGGTGGTGGCCCTTGACTG.TGGAGCTCATGGGCACCGTGGCCAAAC..AGCTGGTGCCAG.....CTGCGTGACTGT
SOM175-e6&7 GTGGTGGTGGCCCTTGACTG.TGGAGCTCATGGGCACCGTGGCCAAAC..AGCTGGTGCCAG.....CTGCGTGACTGT
SOM175-e4 GTGGTGGTGGCCCTTGACTG.TGGAGCTCATGGGCACCGTGGCCAAAC..AGCTGGTGCCAG.....CTGCGTGACTGT

241
VEGF165 GATGCGATGCGGGGGCTGCTGCAATGACGAGGGCCTGGAGTGTGTGCCCACTGAGGAGTCCAACATCACCATGCAGATTA
SOM175 GCAGCGCTGTGGTGGCTGCTGCCCTGACGATGGCCTGGAGTGTGTGCCCACTGGGCAGCACCAAGTCCGGATGCAGATCC
SOM175-e6 GCAGCGCTGTGGTGGCTGCTGCCCTGACGATGGCCTGGAGTGTGTGCCCACTGGGCAGCACCAAGTCCGGATGCAGATCC
SOM175-e6&7 GCAGCGCTGTGGTGGCTGCTGCCCTGACGATGGCCTGGAGTGTGTGCCCACTGGGCAGCACCAAGTCCGGATGCAGATCC
SOM175-e4 GCAGCGCTGTGGTGGCTGCTGCCCTGACGATGGCCTGGAGTGTGTGCCCACTGGGCAGCACCAAGTCCGGATGCAGATCC

321
VEGF165 TC CGGATCAAACCTCACCAAGGCCAGCACATAGGAGAGATGAGCTTCTACAGCACAAACAAATGTGAATGCAGACC...A
SOM175 TCATGATCCGG...TACCCGAGCAGTCAGCTGGGGGAGATGTCCCTGGAAGAACACAGCCAGTGTGAATGCAGACCTAAA
SOM175-e6 TCATGATCCGG...TACCCGAGCAGTCAGCTGGGGGAGATGTCCCTGGAAGAACACAGCCAGTGTGAATGCAGACCTAAA
SOM175-e6&7 TCATGATCCGG...TACCCGAGCAGTCAGCTGGGGGAGATGTCCCTGGAAGAACACAGCCAGTGTGAATGCAGACCTAAA
SOM175-e4 .....CCTAAA

401
VEGF165 AAGAAAGATAG.....AGCAAGACAAGAA...AATCCCTGTGG.....GCCTTGCTCAGAGCGGAGA
SOM175 AAAAAGGACAGTGTGTGAAGCCAGACAGGGCTGCCACTCCCCACCACCGTCCCCAGCCCCGTTCTGTTCCGGGGCTGGGA
SOM175-e6 AAAAAGGACAGTGTGTGAAGCCAGATAG.....
SOM175-e6&7 AAAAAGGACAGTGTGTGAAGCCAGATAG.....
SOM175-e4 AAAAAGGACAGTGTGTGAAGCCAGACAGGGCTGCCACTCCCCACCACCGTCCCCAGCCCCGTTCTGTTCCGGGGCTGGGA

481
VEGF165 .....AAGCA.....TTTGTT....TGAC..A
SOM175 CTCTGCCCCCGGAGCACCTCCCCAGCTGACATCACCCATCCCCTCCAGCCCCAGGCCCCCTCTGCCCCAGCTGCACCCA
SOM175-e6 .....CCCCAGGCCCCCTCTGCCCCAGCTGCACCCA
SOM175-e6&7 .....
SOM175-e4 CTCTGCCCCCGGAGCACCTCCCCAGCTGACATCACCCATCCCCTCCAGCCCCAGGCCCCCTCTGCCCCAGCTGCACCCA

561
VEGF165 A.....GATCCGACAGCTGTAAATGTTCTGCAAAAAC.ACAGACTCG..CGTTGCAAGGCGAGGCAGC
SOM175 GCACCACCAGCGCCCTGACCCCCGGACCTGCCGCTGCCGCTGCCGACGCCGAGCTTCTCCGTTGCCAAGGGCGGGGCT
SOM175-e6 GCACCACCAGCGCCCTGACCCCCGGACCTGCCGCTGCCGCTGCCGACGCCGAGCTTCTCCGTTGCCAAGGGCGGGGCT
SOM175-e6&7 .....
SOM175-e4 GCACCACCAGCGCCCTGACCCCCGGACCTGCCGCTGCCGCTGCCGACGCCGAGCTTCTCCGTTGCCAAGGGCGGGGCT

641
VEGF165 TTGAGTTAAACGAACGTACTTGCAGATGTGACAAGCCGAGGCGGTGA
SOM175 TAGAGCTCAACCCAGACACCTGCAGGTGCCGGAAGCTGCGAAGGTGA
SOM175-e6 TAGAGCTCAACCCAGACACCTGCAGGTGCCGGAAGCTGCGAAGGTGA
SOM175-e6&7 .....GTGCCGGAAGCTGCGAAGGTGA
SOM175-e4 TAGAGCTCAACCCAGACACCTGCAGGTGCCGGAAGCTGCGAAGGTGA

687
```

# FIGURE 6

VEGF <sub>165</sub>	MNFI LSWVHWSLALLLYLHHAKWSQAAPMAEGGGONHHE VVKFMDVYORSYGHIEITLVD	60
SOM175 <sub>short</sub>	MSPI LRRLL LAAALLQAPAO...APVSQPDAPGHORKVSWIDVYTRATQQRREVVP	55
VEGF <sub>165</sub>	IFQIYPDEIEYIFKPSGVPLMRGGGCONDEGLECVPT EESNITMQIMRIKPHOGOHIGEMS	121
SOM175 <sub>short</sub>	LTVELMGTVAKQLVPSGVTVORGGGCPDDGLECVPTGOHOVRMQLMIR.YPSSQLIGEMS	115
VEGF <sub>165</sub>	FLOHINKCECRPKK.....DRA.....ROENPCGSCSERRKHLF.VODPOT	170
SOM175 <sub>short</sub>	LEEHSQCECRPKKKDSAVKPDRAATPHHRPOPRSVPGWDSAPGAPSPADITHPTAPGGLG	175
VEGF <sub>165</sub>	CKCCKNTDSRCKAROLELNERTCRCDKPRR	191
SOM175 <sub>short</sub>	HAAPSTTSALTPGPAAAAADAAASSVAKGGA	207

or...

VEGF <sub>165</sub>	MNFI LSWVHWSLALLLYLHHAKWSQAAPMAEGGGONHHE VVKFMDVYORSYGHIEITLVD	60
SOM175 <sub>long</sub>	MSPI LRRLL LAAALLQAPAO...APVSQPDAPGHORKVSWIDVYTRATQQRREVVP	55
VEGF <sub>165</sub>	IFQIYPDEIEYIFKPSGVPLMRGGGCONDEGLECVPT EESNITMQIMRIKPHOGOHIGEMS	121
SOM175 <sub>long</sub>	LTVELMGTVAKQLVPSGVTVORGGGCPDDGLECVPTGOHOVRMQLMIR.YPSSQLIGEMS	115
VEGF <sub>165</sub>	FLOHINKCECRPKK.....DRA.....ROENP.....G	170
SOM175 <sub>long</sub>	LEEHSQCECRPKKKDSAVKPDRAATPHHRPOPRSVPGWDSAPGAPSPADITHPTAPGGLG	177
VEGF <sub>165</sub>	PGCSERRKHLFVODPOTCKCCKNTDS RCKAROLELNERTCRCDKPRR	191
SOM175 <sub>long</sub>	PRGTOHHOR...PDPRTCRGCRRRSFLRCGGRGLELNPDICGRCKLRR	222

Areas of 100% homology are boxed and conserved residues thought to be involved in homodimerisation are underlined. The VEGF sequence depicted includes the 26 amino acid leader sequence (removal of which gives rise to mature VEGF<sub>165</sub>) giving a total length of 191 amino acids.

Homology of SOM175 to VEGF<sub>165</sub> is 27% (33%) at the protein level, however within this are blocks of 100% homology. In particular, many structural residues are conserved including those thought to be involved in homodimerisation of VEGF (by comparison with PDGF).

*ie.* Cysteine-47

Proline-70, Cysteine-72, Valine-74

Arginine-77, Cysteine-78, Glycine-80, Cysteines-81 & 82

Cysteine-89, Proline-91

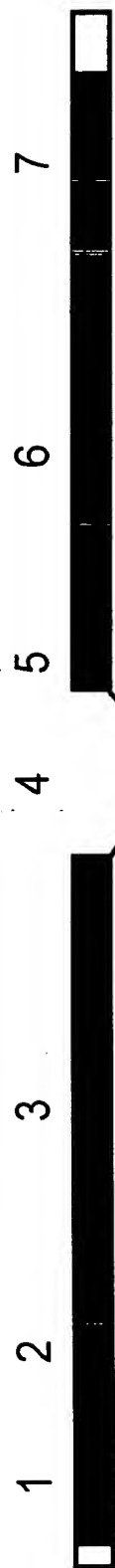
Cysteines 122 & 124

# Splice variants of SOM175

FIGURE 7



SOM175



SOM175 minus exon 4



SOM175 minus exon 6



SOM175 minus exons 6 & 7

FIGURE 8A

Genomic structure of Human SOM175

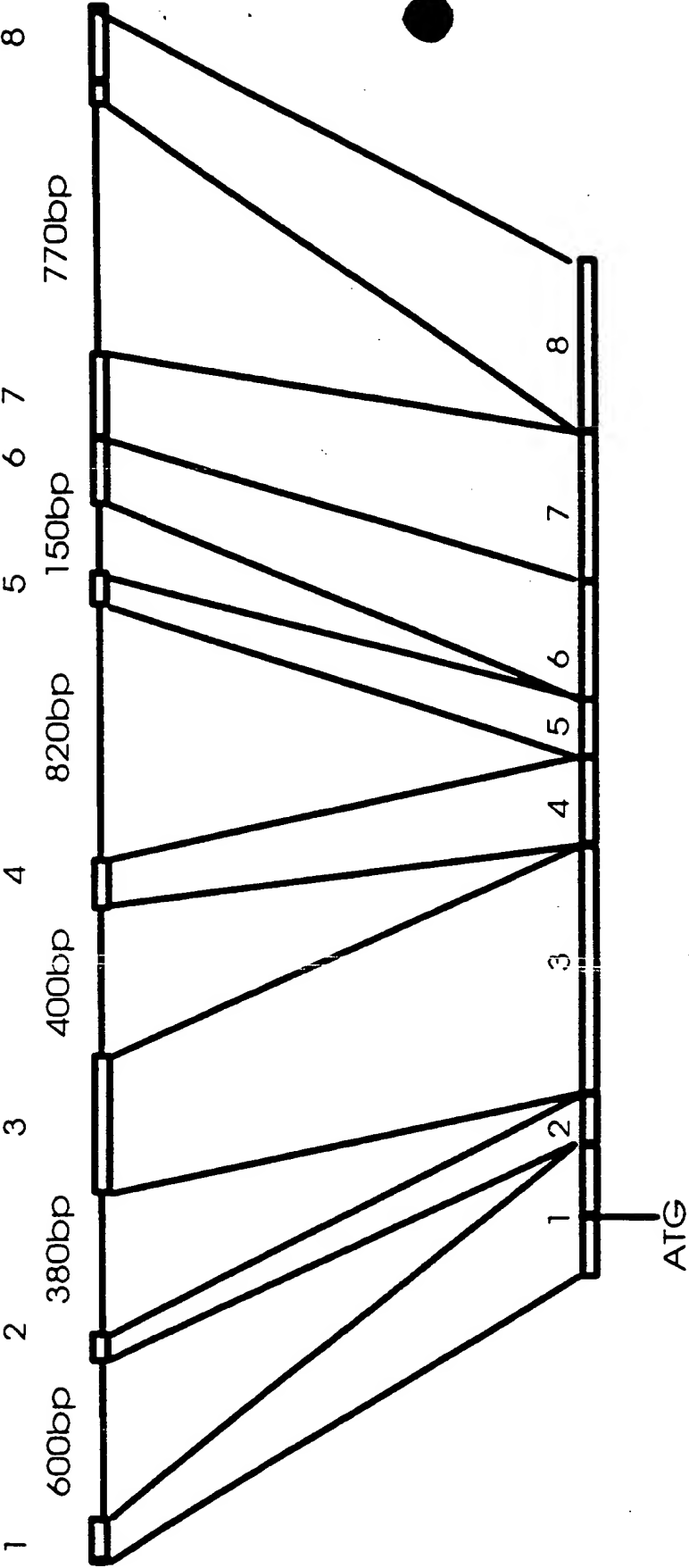


FIGURE 8B

5'UTR... ATGAGG	*Exon 1 (60bp)	GGCCAG gtacgtgagg
tctccacag GCCCCT	Exon 2 (43bp)	GGAAAG aatacttaca
tctgctccca TGGTGT	Exon 3 (187bp)	ATGCAG gtccgagctg
ctgaatacag ATCCTC	Exon 4 (73bp)	ATGCAG gtgtcaggca
acttttcaag ACCTAA	Exon 5 (34bp)	AGACAG gtgagtcttt
ctcctccgta GGCTGC	Exon 6 (101bp)	CTCCAG ccccaggccc
cccactccag CCCCAG	Exon 7 (109bp)	ACCCAG acacctgtag
ccctgctcag GTGCCG	*Exon 8 (22bp)	AGGTGA ...3'UTR

**THIS PAGE BLANK (USPTO)**